

REMARKS

Claims 1-23 are pending. In this paper, claim 2 has been amended to clarify features of the invention. Applicants respectfully submit that the amendments presented herein raise no new issues requiring further searching or consideration by the Examiner.

Reconsideration of the application is respectfully requested for the following reasons.

Claims 1-23 were rejected under 35 U.S.C. §103(a) for being obvious in view of a Gibbon-Nelson combination. This rejection is respectfully traversed for the following reasons.

In finally rejecting claim 1, the Examiner maintained that the Gibbon patent discloses the features of “extracting a plurality of text areas from a video stream” and “calculating importance measures according to weights for each of the extracted text areas.” In support of this position, the Examiner relied on the disclosure at column 8, line 45 - column 14, line 40, of Gibbon.

A. Extracted Text Areas

Here, Gibbon discloses classifying different parts of a multimedia stream. The multimedia stream includes audio, video, and text, where the text is generated from closed captioning or an automatic speech recognition engine. (See column 3, lines 43-45). Unlike claim 1, the text is therefore not included in the original video such as shown, for example, in Figure 5 of Applicants’ drawings. Instead, the closed captioning service superimposes text of a speaker over the video and therefore this text is not included as part of the video. As for the speech recognition engine, this unit converts spoken words into text that never appeared in the video.

Even if the Examiner takes the position that closed-captioned text corresponds to an extracted text area as recited in claim 1, Gibbon makes clear that those closed-captioned areas are not given weights, subsequently selected, and then synthesized into a single synthetic key frame as is further recited in claim 1. (In Applicants' drawings, see, for example, Figure 5 which shows that different portions of displayed text either from the same or different scenes, shots, or frames are combined into a single synthetic frame).

B. Calculation of Importance Measures

Regarding calculation of importance measures, the Gibbon patent discloses that four different classification methods were used for segmenting multimedia content.

One of these methods is the Gaussian Mixture Model (GMM) which the Examiner noted uses weight values, ω_i . However, unlike claim 1, the Gibbon patent discloses using the GMM model only for the purpose of segmenting or separating news from commercials. See column 8, lines 48-49, which provides: "Gaussian Mixture Model (GMM) is employed to model news and commercial classes, individually." The Gibbon patent, therefore, does not teach or suggest "calculating importance measures according to weights for each of the extracted text areas" as recited in claim 1. (Emphasis added).

In terms of applying the GMM model, the Gibbon patent further discloses that news and commercials are separated by the speech of a person talking in the video stream. The Gibbon system determines when a commercial comes on based on differences in speech patterns: "separate news and commercials are identified and segmented based on acoustic characteristics

of audio data.” Columns 5-7 are then devoted to showing how changes in pitch, volume, and frequency are used as a basis for detecting these acoustic characteristics. See also column 9, lines 57, which provides: “The target speaker, background speakers, and other background audio categories are represented by 64 mixture components Gaussian Mixture Models (GMM’s) . . .”

These disclosures further show that any weights generated by the Gibbon system are not assigned to extracted text areas, whether those areas are extracted by the closed-captioned text or the speech recognition engine. See, for example, Figure 5 of Applicants’ drawings which show that actual text depicted in the video are extracted by the claimed invention, weighted, selected, and then synthesized into a single synthetic key frame corresponding to a represented portion of the video.

C. Selection for Synthesis into a Synthetic Key Frame

The Examiner further relied on Figures 13-17 of the Gibbon patent for providing weights to extracted text areas in a video. However, these features only show the results of the output of a text event segmentation unit, which converts the speech of a speaker (e.g., news anchor) shown on the video into text. Unlike the claimed invention, the text itself is not shown in the video, but rather is derived from the speech recognition engine which converts speech from a speaker shown in the video into text. Gibbon does not teach or suggest “synthesizing the number of text areas into a synthetic key frame,” where the number of text areas are “synthesized based upon the importance measures in the order of higher importance” and where

the importance measures are calculated “according to weights for each of the extracted text areas.”

To make up for the deficiencies of Gibbon, the Nelson patent was again relied on.

The Nelson patent discloses a system and method for indexing multimedia documents and creating multimedia queries that include text, image, video, and audio. In the Final Office Action, the Examiner indicated that this system would have been an obvious variant of the claimed invention. However, the Examiner has not indicated where the following features of claim 1 (missing from the Gibbon patent) are taught or suggested in Nelson: “calculating importance measures according to weights for each of the extracted text areas,” “selecting a number of text areas to be synthesized based on the importance measures in the order of higher importance” and then “synthesizing the number of text areas into a synthetic key frame.”

Absent a teaching or suggestion of these features, it is respectfully submitted that a Gibbon-Nelson combination cannot render claim 1 or any of its dependent claims obvious.

Claim 2 separately recites that “the text areas are included in original content of the video in the video stream and are extracted according to certain intervals of the video stream, the text areas in the original content of the video being different from closed-captioned text or text generated by a speech recognition engine.” The Gibbon and Nelson patents do not teach or suggest these features.

Claim 6 separately recites that the “weights are determined in proportion to a size of the text area, a mean text size of the text area and a display segmentation of a text.” These features are not taught or suggested by the cited references, whether taken alone or in combination.

Claim 9 recites that the “weight increases as the size of the text area, the mean text size in the text area or the display duration time of the text increases.” These features are not taught or suggested by the cited references whether taken alone or in combination.

Claim 13 recites features similar to those which patentably distinguish claim 1 from the cited combination. For example, claim 13 recites “calculating importance measures of the text areas by applying the weights”; and “selecting a number of text areas to be synthesized based on the importance measures in the order of higher importance.” These features are not taught or suggested by the Gibbon or Nelson patents, whether taken alone or in combination. Claim 13 further recites synthesizing the text areas into a synthetic key frame. These features are also not taught or suggested. For at least these reasons, it is respectfully submitted that claim 13 and its dependent claims are in condition for allowance.

Claim 17 recites “adding values obtained by multiplying the weight determining factors with corresponding weights to calculate an importance measures for the extracted text areas.” These features are not taught or suggested by the cited references, whether taken alone or in combination. For example, neither reference teaches or suggests calculating importance measures, let alone doing so by adding values obtained by multiplying the weight determining factors with corresponding weights. Absent a teaching or suggestion of these features, it is respectfully submitted that claim 17 and its dependent claims are allowable.

Claims 21-23 were rejected under 35 U.S.C. §103(a) for being obvious in view of a Gibbon-Nelson-Maybury combination. This rejection is traversed on grounds that the Maybury patent publication does not teach or suggest the features of base claims 1, 13, and 17 from which claims 21-23 respectively depend.

In view of the foregoing amendments and remarks, it is respectfully submitted that this application is in condition for allowance. Favorable consideration and allowance of the application are respectfully requested.

To the extent necessary, a petition for an extension of time under 37 C.F.R. 1.136 is hereby made. Please charge any shortage in fees due in connection with the filing of this, concurrent and future replies, including extension of time fees, to Deposit Account 16-0607 and please credit any excess fees to such deposit account.

Respectfully submitted,
FLESHNER & KIM, LLP



Daniel Y.J. Kim, Esq.
Registration No. 36,186

Samuel W. Ntiros, Esq.
Registration No. 39,318

P.O. Box 221200
Chantilly, Virginia 20153-1200
(703) 766-3701 DYK/SWN:knh
Date: AUGUST 8, 2006

Please direct all correspondence to Customer Number 34610